2022, September 8th

Many thanks for your questions and comments on our manuscript, "Stride-TCN for energy consumption forecasting and its optimization". From there, we have made adjustments to improve our article.

We have revised the paper and presented the information in italics, following each reviewer's questions. Changes made in the manuscript are marked using track changes, and we also take notes with each revision briefly on the right side. The revision has been developed in consultation with all coauthors, and each author has approved the final form of this revision. We appreciate your consideration.

Sincerely,
Jin Young Kim

Reviewer #3

1. Line 28: There seems to be a typo at the end.

   This typo has been fixed at line 28.

2. Line 33: It's better to rephrase the sentence "… such as machine learning and deep learning". To me, the terminology is not accurate as deep learning is a part of machine learning. Please also consider other places in the main text such as line 76, etc.

   These typos have been fixed at lines 36, 75 and 76.

3. Line 38: "such as recurrent neural networks (RNNs) and convolutional neural 37 networks (CNNs)". Usually, CNN includes a feature extractor pipeline and mainly this (alone) isn't a very good approach for time series forecasting. It is suggested to rephrase this line and add methods specifically related to time-series forecasting such as LSTM, GRU, Transformers, BERT, etc.

   This typo has been fixed at line 36.

4. Line 60: "This paper is constructed as follows." The following paragraph is missing information of section 1. It is suggested either to add details for section 1 or rephrase the previously quoted sentence to "The rest of the paper is organized as follows" etc.

   This typo has been fixed at line 59.

5. It is suggested to follow the same norm for all the web links. It's better to use a footnote for all the web links. For example, Ref. [4] is a weblink, and the footer [1] is also a web link. In some places, the web links are given in the main text i.e., line 230, line 238, line 244.

   Footnote has been used to cite all the web links at lines 29, 239, 247, 251 and 284.

6. Line 96: "robust in terms of model complexity". This part of the sentence is confusing, it's better to write "lower complexity" and "robust performance".

   This typo has been fixed at line 95-96.

7. Line 105: To achieve consistency and avoid confusion to readers, it is suggested to also add round brackets with $Y = y0, \ldots \hat{y}P$. Moreover, add a comma before $\hat{y}P$.

   This equation has been fixed at line 116.

8. Equation 1: Seems to be mathematically inaccurate as hat is missing in $y0 , . . . , y_p$. Either remove this as it seems to be redundant as function mapping is already provided. Or it is suggested change it to simply Y=f(X).

The equation (1) has been fixed.

9. Line 110: "Building function f is the process of learning to find a network f from a set of time 110 series". This seems to be erroneous, as the learning process doesn't find a network rather the learning process learns the parameters as mentioned in line 88.

   This error has been fixed at line 121.

10. Line 111: the time series X declared doesn't align with X represented in line 105. The X in line 111 seems to be the form of a Set. It is suggested to use common notation throughout the manuscript i.e., you may also covert X in line 105 to a set notation (use curly brackets).

    This error has been fixed at line 114.

11. Line 114: Moreover, add a comma before $\hat{y}$P.

    This error has been fixed at line 123.

12. There is a miss-aligned ":" on line 122.

    This error has been fixed at line 132.

13. In Equation 2: I am a bit confused here, what is "y" is this the same "Y" mentioned in line 105? Moreover, the loss function notation i.e., L_delta must be aligned as mentioned in line 114.

    "Y" *in equation 2 and* "Y" *mentioned in line 116 are the same, the equation (2) has been fixed by using uppercase Y.*

14. In Equation 3: Please avoid using a dot product notation i.e., "·", as it may confuse the readers. It's better to use multiplication or no symbol at all. Moreover, what are x and f? Additionally, it's better to change the notation of f as it has already been used to represent a function.

    We have changed the notation f *to* w *to avoid ambiguity. Therefore,* x *is a sequence input and* w *is a filter. The change has been updated from lines 148 to 152 and modified Equation 3.*

15. Line 149: The notation d doesn't align with Figure 1. Please make it consistent throughout the manuscript.

    d has been used consistently in Figure 1 and throughout the manuscript, lines 161-162.

16. Figure 1: What is the end range/value of L?

    As explained in line 161, in TCN, $L = \{1,2,4,\dots,2^N\}$ *where* N *is the number of layers.*

17. Please use a consistent representation of K in Equation 3, Table 1, Table 2, Line 140, etc.

    We have revised and used a consistent representation of K in equation (3), Table 1, 2 and Line 152.

18. Figure 2: It would be more convenient if the time scale in terms of weeks, months, or years is mentioned. It will provide a better picture to readers.

    We have edited Figure 2 to provide better intuition to readers.

19. Line 174: "The Ln layers are calculated directly based". This sentence is not clear. Do you mean the layer calculation or layer parameters?

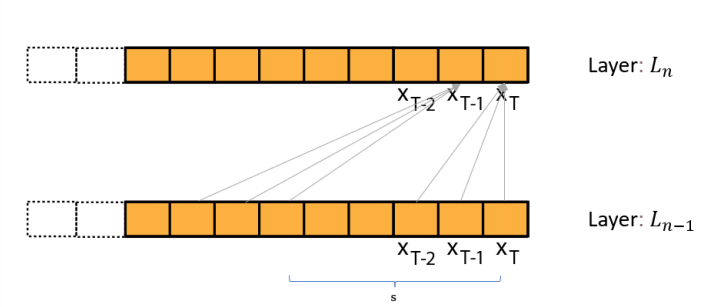

Figure 1

L_n is n^{th} layer used in calculation. We have revised this information in lines 186-187.

20. Algorithm 1, Algorithm 2: Some notations should be explained as the notations are not standard, i.e., both arrows, := . Moreover, try to use consistent loop notation for both algorithms.

    We have edited Algorithms 1 and 2.

21. Line 201, 225, 248: There seems to be a problem with the reference.

    We have corrected the references at lines 213, 236 and 257.

22. Algorithm 2, Line 4: Please mention the loss notation as described earlier in the manuscript.

    We have mentioned the loss notation as described earlier in Algorithm 2.

23. Equations 6 and 7: These Equations are conflicting with y0 , . . . , yp. It creates confusion about whether y_i in these equations is an element or a vector.

    We have capitalized the notation "Y" to avoid ambiguity, and y in equations 6 and 7 is an element.

24. In line 287: If the loss function for training is MSE, then why Huber loss is introduced on line 121?

In this study, we have employed Huber loss as an objective function to train TCN models, while MSE is for training LSTMs and GRUs. The decision to use different objective functions here is based on our experimental results and prior research evidence. For fair comparisons, However, we only use MSE as a metric to evaluate the models. The explanation has been added to lines 282 and 294.

25. Line 318: The mentioned Figure 3 doesn't indicate the error rates. It seems that the figure reference is inaccurate. Please reconfirm all the figure references.

    We have revised and fixed the 'Figure 3' to 'Table 5' references at line 326.

26. Please remove the line "In summary" from the conclusion.

    We have removed this line, the change can be checked at line 387.